

Price Benchmark Regulation of Multiproduct Firms: An Application to the Rail Industry

Wesley W. Wilson *University of Oregon*

Frank A. Wolak *Stanford University*

Abstract

Building on Harold Demsetz’s argument that market forces can be more effective in disciplining the pricing behavior of firms relative to explicit output price regulation, this paper proposes and analyzes the performance of a price benchmark approach to identify unreasonable prices in multiproduct industries subject to residual reasonable price regulation. An empirical application to the freight rail sector demonstrates that this mechanism has the potential to provide economically meaningful relief to shippers at lower cost, with less administrative burden, and without significant adverse financial consequences for the railroads.

1. Introduction

Harold Demsetz disputed the classical justification for output price regulation of natural monopolies (Demsetz 1968). He argued that, absent legal barriers to entry, firms would ultimately price below the monopoly level because of the threat or realization of competitive entry. Demsetz was also a critic of what he called the nirvana approach to public policy economics, which “presents the relevant choice as between an ideal norm and an existing ‘imperfect’ institutional arrangement” (Demsetz 1969, p. 1). He favored “a *comparative institution* approach in which the relevant choice is between alternative real institutional arrangements”

This paper evolved from the authors’ participation in a 2015 National Academy of Sciences panel review of regulation in the railroad industry, “Modernizing Freight Rail Regulation.” The other panel members—Ken Boyer, Jerry Ellig, José A. Gómez-Ibáñez, Anne Goodchild, Thomas Menzies, and Dick Schmalensee—provided invaluable feedback throughout the process of completing this paper. Earlier versions were presented to the Association of Transport, Trade, and Service Studies; the Railroad-Shipper Transportation Advisory Council; the Advanced Workshop in Regulation and Competition; and the Transportation and Public Utilities Group at the 2019 Allied Social Science Association meetings. The authors gratefully acknowledge comments and discussions that emanated from those audiences. We would also like to thank Ray Atkins, Jerry Ellig, William Huneke, Patrick McCarthy, Tom Menzies, Keaton Miller, Russell Pittman, David Sappington, Dick Schmalensee, and Clifford Winston for detailed comments on a previous draft. Sam Peltzman and a referee provided helpful suggestions. We thank Kevin Henrickson and Rodrigo Naumann for research assistance.

[*Journal of Law and Economics*, vol. 65 (February 2022)]

© 2022 by The University of Chicago. All rights reserved. 0022-2186/2022/6501-0013\$10.00

(Demsetz 1969, p. 1). This paper builds on those two insights to propose methodology for protecting consumers from excessive prices in multiproduct industries.

A number of formerly regulated multiproduct industries have a transitional or permanent residual regulatory mandate to protect consumers from excessive prices. The Federal Energy Regulatory Commission is required to ensure that all wholesale electricity prices are just and reasonable and not unduly discriminatory or preferential, even in parts of the United States with formal offer-based short-term markets for wholesale electricity. The commission has a similar regulatory mandate for natural gas and oil transportation despite the fact that prices for most natural gas and oil movements are set through bilateral negotiations between the pipeline owner and the purchaser of wholesale natural gas or oil. In the aftermath of the Airline Deregulation Act of 1978 (Pub. L. 95-504, 92 Stat. 1705), the Civil Aeronautics Board had a transitional mandate to ensure that airfares were not unjust and unreasonable. The Staggers Rail Act of 1980 (Pub. L. No. 96-448, 94 Stat. 1895), which partially deregulated the railroad industry, imposes a statutory mandate on the Surface Transportation Board (STB), the industry regulator that replaced the Interstate Commerce Commission (ICC) to protect captive shippers from excessive prices.

These regulatory mandates have proven challenging to enforce to the satisfaction of the parties involved because of the conceptual difficulty in defining a reasonable price for a multiproduct firm with substantial economies to scope and scale in production. Darius Gaskins (2008, p. 561), the former chief executive officer of Burlington Northern Railroad, argues that this residual regulatory challenge in the railroad industry “still has not been solved to everyone’s satisfaction after 150 years of effort.” That is because railroads provide thousands of products, depending on the commodity and distance shipped, and the incremental cost of a shipment and the marginal cost of including an additional ton in the shipment exclude the vast majority of the railroad’s total cost of production.¹ This implies that setting each shipment price equal to either the average incremental cost of the shipment or the marginal cost of shipping an additional ton of that product would not provide sufficient revenue for the railroad to cover its fixed and common costs. Prices above the average incremental cost and the marginal cost of shipping an additional ton for a substantial fraction, if not all, of its shipments are necessary for the railroad to recover its total cost of production.

This regulatory challenge falls squarely in the realm of Demsetz’s comparative institution approach because the information requirements on production technologies, input costs, and consumer demands necessary to implement the efficient solution are prohibitive.² Instead, the relevant economic policy question

¹ For a rigorous definition of the incremental cost of a shipment and the marginal cost of including an additional ton in a shipment, see Wilson and Wolak (2016). Costs not caused by a shipment or by moving an additional ton in a shipment are excluded from these measures. These include the cost of the track, rolling stock, management salaries and benefits, and the vast majority of labor costs.

² Laffont and Tirole (1990a, 1990b) present models of optimal regulation of multiproduct firms in monopoly and competitive environments that illustrate these information requirements.

is, in the language of Demsetz (1969, p. 1), which “alternative real institutional arrangement seems best able to cope with the economic problem”? Currently, the STB enforces the prohibition against excessive prices in the Staggers Rail Act through a two-step administrative process that must first establish that the railroad is dominant in the provision of the shipment. If the railroad is found to be dominant, the STB must then determine if the tariff charged is excessive. As discussed in Pitman (2010), TRB-NAS (2015), and Rate Reform Task Force (2019), the administrative process used to make this decision can be costly, which may have limited the number of rate relief cases filed, particularly by small shippers. Between 1996, when the STB came into existence, and 2021, there have been only 52 cases filed, and an unreasonable rate was found in only 12 of them.³ Virtually all of these rate relief cases were brought by large shippers of coal and chemical products, who may have been better able to rationalize the cost of participating in the STB rate relief process.

These features of the STB rate relief process have increasingly led to calls for its reform, particularly by small shippers that would like a low-cost mechanism for obtaining rate relief. To this end, we propose a price benchmark approach to establish whether a railroad is dominant in provision of a shipment and final-offer arbitration between the two parties to determine whether the tariff charged is excessive. We also describe how our price benchmark can be used in final-offer arbitration. We present theoretical and empirical evidence that our proposed rate relief process is a superior comparative institution solution to the economic problem addressed by the STB’s current approach.

A major challenge to meaningful reform of the STB rate relief process is that traditional cost-based approaches to price regulation are not available for a multiproduct firm with pricing flexibility for a significant fraction of the products it sells.⁴ The large number of products sold by railroads and the large share of common costs in a railroad’s total cost of production imply that even if the railroad’s multiproduct cost function is known with certainty, this would not make the job of determining whether a shipment price is excessive any easier. The STB would be able to compute the average incremental cost of a shipment or the marginal cost of shipping an additional ton with certainty, but this would merely change the STB’s problem from determining whether a price is excessive to the equally difficult problem of determining whether the markup over the average incremental cost or marginal cost of the shipment is excessive.⁵

³ Of the remaining 39 cases, two were withdrawn, 27 resulted in a settlement between the railroad and shipper, and in 11 the tariff rate charged was ultimately found to be reasonable.

⁴ The regulator can set the prices of products only if the firm does not have pricing flexibility for these products, which still leaves open the question of what fraction of fixed and common costs the firm is entitled to recover from these products.

⁵ Further complicating the process of crafting an alternative rate relief process is the fact that there are unlikely to be significant short-run total welfare losses associated with railroads charging shippers excessive prices. A profit-maximizing railroad would not charge a shipper a price that is so high that a movement with an economic benefit to the shipper greater than its incremental cost to the railroad would not occur. Setting the shipment price high enough to curtail the shipment would reduce total contributions to the recovery of the railroad’s fixed and common costs. Nevertheless, the

This paper analyzes an alternative approach to making this determination that does not require knowledge of the multiproduct firm's cost function or the structure of the demand for products sold by the firm. Our benchmark price approach could be part of a low-cost mechanism for captive shippers to obtain rate relief by replacing the current approach to determining whether a railroad is dominant for a shipment and also by using it to provide input about whether a rate charged by a dominant railroad is excessive. This methodology relies on circumstances that are increasingly prevalent in many formerly regulated industries: the existence of a large sample of reasonable prices for products and the observable characteristics of each product. We use this sample to estimate nonparametrically the conditional distribution of reasonable prices given a vector of observable product characteristics and then use this conditional distribution to construct a price benchmark based on the value of the observable characteristics for a shipment suspected of having an excessive price. If the actual price exceeds the price benchmark, then the shipment price would be deemed to be the result of the actions of a dominant railroad and worthy of further regulatory scrutiny to determine whether the price is excessive and the shipper is due rate relief. As we discuss below, this mechanism would be low cost for the STB to implement and for any shipper to access and thereby will increase the opportunities for rate relief for small shippers.

There are two important considerations in setting the value of a price benchmark for a shipment. First is the probability of false positives: reasonable prices that are incorrectly found to exceed the price benchmark. Second is the possibility of false negatives: unreasonable prices that are incorrectly found not to exceed the price benchmark. We investigate this issue with a simulation study in which we first estimate the conditional distribution of reasonable prices given shipment characteristics on data simulated from markets in which the firm faces a reasonable level of competition and then apply our price benchmark methodology using this estimated distribution with data simulated from a mixture of reasonable and unreasonable shipment prices.

Similar to the case of statistical hypothesis testing, a rule for setting the price benchmark that minimizes the sum of squares of misclassification errors argues in favor of an approach that requires overwhelming statistical evidence against a price being reasonable before it is deemed unreasonable. Our simulation results find that setting the value of the price benchmark between the upper fifth percentile and upper first percentile of the conditional distribution of reasonable prices given the shipment characteristics minimizes the sum of squared misclassification errors for a range of plausible distributions of unreasonable shipment prices.

As noted above, our benchmark price approach can also provide valuable input

railroad is likely to charge a captive shipper a rate that transfers virtually all of the economic surplus associated with the movement to the railroad. In the longer run, the possibility of excessive pricing is likely to deter future investments at the site and at other sites unless perfectly efficient long-term contracts can be written. With uncertainty and asymmetric information, such perfect contracts cannot be executed. Hence, there is a reason to limit excessive pricing.

to the process used to determine whether the price charged is excessive and what price should be set if the current price is found to be excessive. Consequently, another important consideration in the design of our approach is the revenue impacts to the railroad of resetting shipment prices that violate the price benchmark to a mitigated value at or below the price benchmark. If violations of the price benchmark occur too frequently and if the mitigated shipment price is set too low, there is a risk that a railroad that is revenue adequate—earning sufficient revenues to recover its total cost of production—may become revenue inadequate. Using the choice of the price benchmark recommended by our simulation study, we explore the impact of different choices for the mitigated or reasonable shipment price in the event that an actual price violates our price benchmark with data from the STB's Carload Waybill Sample (CWS) for four broad classes of shipments: petroleum products, farm products, coal, and chemical products. In all cases, we find that for our choice of the price benchmark, resetting the price to any of our three choices for a reasonable price for the shipment has a very small percentage impact on the total revenues earned by the railroads shipping these products.⁶ Despite these small aggregate revenue impacts, for all of our choices of the reasonable price, we also find that the average value of the shipment-level difference between the actual price and the reasonable price is a substantial percentage of the average value of the actual price of these shipments. Consequently, this use of our price benchmark procedure also provides significant rate relief to a shipper facing prices that are determined by our price benchmark to be set by a dominant railroad.

Another concern with our price benchmark approach is the credibility of the estimate of the price benchmark for all shipment characteristics in the set of potentially unreasonably priced shipments. We explore this hypothesis by dividing the set of potentially unreasonably priced shipments into those with shipment characteristics that lie in the convex hull of the vectors of shipment characteristics in our reasonably priced sample of shipments. Because the asymptotic variance of our estimate of the price benchmark for a given vector of shipment characteristics is decreasing in the value of the density of the vector of shipment characteristics at that point, more precise estimates of the price benchmark are likely for values of the vector in the convex hull of shipment characteristics for the reasonably priced sample (for the proof, see Li and Racine 2007, theorem 2.2). For three of the four broad classes of shipments we consider, the frequency of violations of our benchmark price, the average price reductions, and the percentage of changes in railroad revenue for each approach to price mitigation for this restricted sample of potentially unreasonably priced shipments is not significantly different from the corresponding values for our full sample of potentially unreasonably priced shipments. For the remaining class of products, coal, a significant number of shipments have a vector of characteristics that lie outside the convex

⁶ A less conservative approach to setting the value of the price benchmark at, for example, the conditional median price for the observed shipment characteristics leads to larger revenue losses relative to resetting the actual price to the price benchmark level.

hull of the vectors of shipment characteristics in our reasonably priced sample of shipments. However, for this restricted sample, the frequency of violations of our price benchmark, average price reductions, and percentage of changes in railroad revenue for each approach to price mitigation are not significantly different from the corresponding values for our full sample of potentially unreasonably priced shipments.

Finally, we provide recommendations for how our price benchmark approach can be used to provide rate relief for captive shippers in a manner that is consistent with the current two-step STB process. The first step of establishing that a shipment price is the result of a dominant railroad would be answered in the affirmative if the actual shipment price exceeded our price benchmark for that shipment. The STB could consider other factors in making this finding, but relying primarily on our price benchmark approach has the advantage of simplicity and consistency with the economic logic described above. The next step in the process would be to determine whether the rate charged is excessive. We believe that this should be accomplished through final-offer arbitration. Our empirical results for the revenue effects on the railroad and rate relief for shippers from using results from our price benchmark approach suggest that it can provide useful input to this arbitration process as well.

Section 2 summarizes the pre- and post-Staggers Rail Act regulatory framework governing the railroad industry. It also summarizes the shortcomings of the current approach to regulating unreasonable prices charged to captive shippers and why we believe our benchmark price approach overcomes many of these shortcomings. Section 3 outlines our approach to estimating the conditional distribution of reasonable prices given shipment characteristics. Section 4 presents the simulation study we use to compare methodologies for computing the value of the price benchmark from the conditional distribution of reasonable prices given product and shipment characteristics. Section 5 reports the results of applying our methodology to the STB's CWS for four classes of shipments to assess the extent of rate relief obtained by the shipper and the impact of the rate relief on annual railroad revenues for different approaches to determining a reasonable price for a shipment that is found to have an excessive price. Section 6 presents our recommendations for using our benchmark price methodology to carry out the STB's statutory mandate to protect captive shippers. Section 7 summarizes our results and proposes directions for future research.

2. A Brief History of Railroad Regulation

The history of regulatory oversight of the railroad industry since the inception of the ICC in 1887 can be divided into the pre-Staggers Rail Act period, when prices and entry into and exit from the rail sector were regulated by the ICC, and the post-Staggers Rail Act period, when railroad price regulation and entry and exit regulation were largely eliminated. Residual regulation of these functions was conducted by the ICC until 1995, when it was eliminated by the ICC Sunset Act

(Pub. L. 104-88, 109 Stat. 803). The act also created the STB, which now carries out these functions. For both the pre- and post-Staggers Rail Act regimes, we highlight the regulatory challenges facing the railroad industry. This section concludes with a discussion of why we believe the use of our price benchmark should reduce the cost and improve the effectiveness of the STB's current approach to residual price regulation, particularly for small shippers.

2.1. Pre-Staggers Rail Act Regulation

Prior to the passage of the Staggers Rail Act, rates for all railroad shipments were subject to approval by the ICC (for a comprehensive policy survey, see Keeler 1983). Rate proposals were typically provided by rate bureaus composed of railroad staff that operated cooperatively with antitrust immunity. The ICC would then review the rate proposals and frequently prohibited their implementation or significantly reduced them before they were allowed to be implemented. As Stone (1991) notes, rate reductions to respond to competition from other modes of transportation were often blocked. This regulatory structure did not encourage efficient operation of the rail network or maximize the revenues earned by the railroads. By the early 1970s, it was clear that the industry was in extremely poor financial condition, as evidenced by a number of high-profile bankruptcies such as the Penn Central in 1970; the Chicago, Rock Island, and Pacific in 1975; and the Milwaukee Road in 1977. Congress subsequently nationalized the Penn Central (also known as Conrail) and was spending billions of dollars per year to keep the industry afloat. By the late 1970s, Congress tired of underwriting the industry, and it was widely held that regulatory reform was necessary to allow the railroads to meet new forms of competition and foster innovation (on regulatory reform, see Keeler 1983; Gallamore and Meyer 2014).

The Railroad Revitalization and Regulatory Reform Act of 1976 (4R Act; Pub. L. No. 94-210, 90 Stat. 31) and the Staggers Rail Act of 1980 provided for significantly reduced federal regulatory oversight of the railroads. This was accomplished by introducing new mechanisms governing the regulation of rates, allowing confidential contracts between railroads and shippers at negotiated rates, and easing impediments to rail line abandonment and mergers.

2.2. Post-Staggers Rail Act Regulation

The 4R Act and the Staggers Rail Act placed a greater emphasis on market forces to discipline rates. The effects of these legislative changes on the railroad industry have been dramatic, with substantial decreases in costs, rates, and the size of the rail network and a tremendous consolidation of firms (for examples, see McFarland 1989; Barnekov and Kleit 1990; Berndt et al. 1993; Velturo et al. 1992; Burton 1993; Wilson 1994, 1997; MacDonald and Cavalluzzo 1996; Grimm and Winston 2000; Ellig 2002; Bitzan and Keeler 2003; Government Accountability Office 2006; Bitzan and Wilson 2007; Winston et al. 1990; Schmalensee and Wilson 2016).

The legislation anticipated the need to protect shippers that do not have an economically viable alternative for a shipment. It established the notion of market dominance to protect these so-called captive shippers from excessive rates. The STB had the jurisdiction to consider the reasonableness of a rate only if there was a finding that the railroad was market dominant over the movement.

Market dominance is defined as the absence of effective competition from other railroads or modes of transportation (49 U.S.C. sec. 10707). A rate for a shipment is automatically considered reasonable if the revenue the railroad receives (R) for it does not exceed 180 percent of the railroad's "variable cost" (VC), as determined by the STB's costing methodology (49 U.S.C. sec. 10707(d)(1)(A)).⁷ If a disputed rate fails the $R/VC \leq 180$ test and is found to be serving a market lacking effective competition, the railroad is deemed to be market dominant, and the STB can rule on whether the rate is excessive.⁸ If the STB subsequently makes this finding, it must order the railroad to compensate the shipper for overpayments, and it may prescribe the maximum rate the railroad can charge for future movements (49 U.S.C. secs. 11704[b], 10704[a][1]).

Shippers historically brought excessive-rate cases under the stand-alone cost (SAC) criteria introduced by the ICC in 1985 (Coal Rate Guidelines, Nationwide, 1985 [1 I.C.C.2d 520, 1985 WL 56819]).⁹ In this type of case, the SAC of a hypothetical railroad providing the shipment is used to establish an upper bound on the rate that is deemed reasonable for the shipment. The time and effort required to make an SAC claim against a railroad are substantial. The process of determining the SAC for a hypothetical railroad providing the shipment is extremely complex, with ample room for disagreement between parties about the many assumptions underlying the calculation. Each of these points of disagreement must be litigated at the STB, which makes the process both expensive and time-consuming. The STB estimates that the costs of pursuing an SAC case can exceed \$5 million (Rate Regulation Reforms, Ex Parte No. 715, pp. 10–11 [STB, July 8, 2013]).

The cost and complexity of SAC cases have led to a number of legislative and policy changes by the STB to reduce the time and cost of filing for rate relief. In the ICC Termination Act of 1995 (Pub. L. No. 104-88, 109 Stat. 803), Con-

⁷ We put "variable cost" in quotation marks to denote the fact that this is primarily an allocated-cost-based regulatory construct that is inconsistent with the economic theory of multiproduct cost functions for the reasons given in Wilson and Wolak (2016).

⁸ See Eaton and Center (1985) and Wilson (1996) for more details on the process used to determine market dominance. Until recently, excessive-rate cases had to be initiated by the shipper, but the Surface Transportation Board (STB) Reauthorization Act of 2015 now authorizes the STB to investigate on its own initiative (Pub. L. No. 114-110, 129 Stat. 2228, sec. 11).

⁹ In the Coal Rate Guidelines, the commission adopted the principle of constrained market pricing wherein a captive shipper should not be required to pay more than necessary for the rail carrier to earn adequate revenue for the service provided or for any productive inefficiencies in how the rail carrier provides the service. Hence, early excessive-rate cases could be brought under three standards: the Stand-Alone Cost (SAC) test, excessive profits earned by a railroad that was revenue adequate, and management inefficiencies included in the rate charged. However, all cases were brought under the SAC criteria until the simplified guidelines for bringing rate cases were introduced in 1996.

gress ordered the STB to develop expedited procedures for resolving disputes. In response, the STB introduced the Three-Benchmark standard in 1996. In 2006, the STB revised the Three-Benchmark rules and introduced the Simplified SAC rules. Both of these expedited procedures limit the evidence that parties can submit and set a time limit for decisions (for more complete discussions, see Pittman 2010; Wilson and Wolak 2016). However, the simplified procedures also limit the amount of refunds a shipper can obtain from excessive prices.

To implement the first stage of the excessive-rate test, the VC of the shipment under consideration must be calculated, and the legislation mandates that the STB have a costing methodology.¹⁰ To comply, the ICC developed the Uniform Railroad Costing System (URCS), which was adopted in 1989 and shares a methodological approach with earlier allocated-cost accounting schemes used by the ICC.

Elsewhere, we examine the theoretical and empirical validity of the URCS methodology for computing the VC of a shipment (Wilson and Wolak 2016). We argue that it is an ad hoc cost allocation methodology that is inconsistent with the economic theory of multiproduct cost functions and irrelevant to how a profit-maximizing railroad (not subject to rate review on the basis of this shipment cost measure) would set the price of a shipment. There are many instances in the CWS for all product categories of railroads providing a shipment at a price that is less than the URCS VC of the shipment (see Wilson and Wolak 2016), which implies that the railroad is receiving less revenue than it costs to provide the shipment. This irrational behavior by railroads implied by the URCS costing methodology and its inconsistency with the economic theory of multiproduct cost functions argues against the use of the URCS VC in determining excessive rates. We conclude (Wilson and Wolak 2016) that URCS costs do not meet the law's requirement for an economically accurate shipment cost and therefore have little relevance for determining the reasonableness of the price charged for a given unit of traffic, contrary to its use in the law's R/VC formula. The STB's own Railroad-Shipper Transportation Advisory Council (2011, p. 1) refers to the URCS as "an outdated and inadequate costing system."

It is important to emphasize that even if the STB had access to perfect measures of the incremental cost for all possible shipments a railroad could provide, this information would not get it any closer to determining what is an excessive price or what a reasonable price is for a shipment because of the substantial fixed and common costs associated with providing rail service. Financial viability of a railroad requires it to charge prices in excess of the incremental and marginal cost of a shipment for a significant fraction or all shipments in order to recover its fixed and common costs. With perfect estimates of the incremental and marginal costs of a shipment, the STB would face the equally challenging tasks of determining what shipment price is an excessive markup over the average incremental cost

¹⁰ The Staggers Rail Act (sec. 10705a[m][1]) requires the ICC to determine the VCs of a shipment by using its Rail Form A costing method or to adopt an alternative method.

or marginal cost of a shipment and what shipment price is a reasonable markup over those measures.

Our price benchmark approach explicitly addresses these challenges by using information from shipments that the STB has determined are reasonably priced to determine what is an unreasonable price for a shipment. Our price benchmark is significantly less costly for shippers to access in terms of time and legal expense. Unlike the Three-Benchmark approach described above, it does not involve the use of the URCS costing methodology. Our methodology also makes use of the fact that an increasing number of shipments are occurring at negotiated rates or rates where the STB has determined that the railroad faces effective competition for those shipments.

Prices for shipments that the STB has determined are reasonably priced are used to estimate the conditional distribution of shipment prices given the observable characteristics of a shipment that account for differences in shipment costs, the commodity shipped, and other characteristics. We then use this estimated conditional distribution to compute a price benchmark for a shipment (based on its observable characteristics) that has a potentially unreasonable price set by a dominant railroad. An extreme percentile of the reasonable-price conditional distribution is the benchmark relative to which the actual price is compared in order to determine if it is an unreasonable price set by a dominant railroad. This approach can be applied to all markets, railroads, and commodities utilizing data that are easily obtained and/or collected by the STB, primarily through the CWS.

We do not claim that our price benchmark distribution estimation procedure recovers the conditional distribution of competitive prices given shipment characteristics, only that it recovers the conditional distribution of prices for the set of shipments that the STB has determined have a reasonable price. Our empirical application using the CWS data assumes specific criteria for a shipment to enter the estimation sample, but our approach allows the STB to use different criteria for selecting the reasonable-price conditional distribution estimation sample.¹¹ Different criteria for selecting the reasonable-price sample will lead to a different reasonable-price conditional distribution estimate. However, because the benchmark price is the upper α th quantile, for a small value of α , of the conditional distribution, changes in the sample used to estimate the conditional distribution of reasonable prices are likely to have a limited impact on the set of shipments that violate the price benchmark in the potentially unreasonably priced sample.

The assumption behind our approach is that the conditional distribution of prices given a broad class of product and shipment characteristics estimated from the sample of reasonably priced shipments provides a valid estimate of the distribution of reasonable prices given any vector of observed shipment characteristics. The choice of this reasonable-price sample is a decision that should be made

¹¹ We also recognize that railroads may take into account which shipments are used to construct our price benchmark distribution in how they price these shipments. However, the increasing number of shipments that are exempt from STB maximum-price regulation and the methodology we use to compute the price benchmark described in Section 4 argues against this being a significant shortcoming of our approach.

by the STB in consultation with shippers, railroads, and other interested parties. The STB's relative preference for reducing the probability of falsely finding that a reasonable shipment price is unreasonable versus reducing the probability of failing to reject that an unreasonable shipment price is reasonable determines the percentile of the conditional distribution of reasonable prices that becomes the price benchmark for determining if the observed price for the shipment is excessive. As we demonstrate below, this process is very similar to the problem of choosing the size of a statistical hypothesis test for which the researcher must balance the probability of a type I error against the probability of a type II error.

3. The Reasonable-Price Conditional Distribution

This section summarizes our procedure for estimating the distribution of a reasonable price for a shipment conditional on product and shipment characteristics. This conditional distribution is the essential input for computing the price benchmark used to determine whether a shipment price is the result of a market-dominant railroad and is therefore unreasonable. Our approach is completely nonparametric and relies only on the existence of a sample of reasonable shipment prices and shipment characteristics for a product group, such as the one available from the CWS. We then describe the contents of the CWS data and how they are used to estimate the conditional distribution, leaving the technical details of the estimation procedure to the Appendix. Finally, we describe how the estimation process can be automated to update the conditional distribution each year with new data.

3.1. Conditional Distribution Function

Computing the benchmark price requires an estimate of the conditional distribution function of Y , the shipment price, given \mathbf{X} , a J -dimensional vector of product and shipment characteristics. If Y_i is the price of shipment i and \mathbf{X}_i is the vector of observable characteristics of shipment i , then $F(y | \mathbf{X}_i) \equiv \text{Prob}(Y_i \leq y | \mathbf{X}_i)$, the probability that Y_i is less than or equal to y given the vector of shipment characteristics \mathbf{X}_i . Our procedure for estimating $F(y | \mathbf{X})$ does not impose a parametric functional form on the relationship between y and \mathbf{X} , and it can be automatically updated by the STB as new CWS data become available. We use the Nadaraya (1964) and Watson (1964) kernel regression estimator of $F(y | \mathbf{X})$, which accounts for stratified sampling of shipments in the CWS data (on nonparametric models, see Pagan and Ullah 1999). Computing the value of $\hat{F}(y | \mathbf{X})$, our estimate of $F(y | \mathbf{X})$, for a potentially unreasonable shipment price y^* and vector of shipment characteristics \mathbf{X}^* , requires computing $\hat{F}(y^* | \mathbf{X}^*)$ using the process described in the Appendix.

3.2. Data

The CWS is the primary data source used to estimate the conditional distribution of reasonable prices given a shipment's observed characteristics. Each year's CWS consists of more than 500,000 stratified randomly sampled shipments with

information on revenue, distance, shipment size, and the identities of the railroads that provided the service.

The CWS records also contain codes that we linked with the Oak Ridge National Laboratory Rail Network files to allow shipper and receiver locations to be identified. In particular, rail station records are identified by a standard point location code. These identifiers permit mapping of origin and destination stations in the CWS and the assignment of latitude and longitude values to each shipment origin and destination. These data, along with railroad network geographic information system data, are combined to identify the locations of stations and shipment origins and destinations and to develop the measures of railroad competition described below. The data are also used in conjunction with the Port Series data produced by the US Army Corps of Engineers to measure the presence of water competition. The Port Series data indicate the locations of ports on US waterways and the commodities handled by each port.

All rates from the CWS are adjusted to constant 2009 dollar values using the quarterly gross domestic product price deflator available from the Federal Reserve Economic Data through the Federal Reserve Bank of St. Louis.¹² We use a subsample of CWS movements that are exempt from regulatory oversight to estimate the reasonable-price conditional distribution. For the products we consider, this subsample is composed of the two classes of movements created by the Staggers Rail Act: exempted traffic and contract movements.¹³

When the Staggers Rail Act was passed, our approach was not feasible because all rates were subject to regulation. The act allowed the regulatory authority to exempt traffic from maximum-price regulation (49 U.S.C. sec. 10502) and the use of confidential contracts that are not subject to regulation. This new regulatory policy would allow “competition and the demand for services to establish reasonable rates for transportation by rail” (49 U.S.C. sec. 10101[1]). Regulators were instructed to be aggressive in fully exempting from any further regulatory control all traffic—truck-competitive traffic being the most obvious—for which regulation was “not needed to protect shippers from the abuse of market power” (49 U.S.C. sec. 10502).¹⁴ Once it designated a class of traffic to be exempt, the ICC would no longer have control over the rates charged to shippers or the amount and quality of service made available to them. For commodities that were not ruled exempt, a critical reform was the law’s legalization of confidential contracts between railroads and shippers. Any shipment moved under contract would be

¹² See Federal Reserve Bank of St. Louis, Gross Domestic Product: Implicit Price Deflator (<https://fred.stlouisfed.org/series/GDPDEF>).

¹³ Other definitions of reasonably priced shipments could be used. For example, an early version of this methodology reported in TRB-NAS (2015) uses exempted traffic and contract shipments with more than one railroad serving the origin or destination or water transport availability at the origin or destination. This definition of the sample of reasonably priced shipments yields results similar to the ones reported in Section 5.

¹⁴ Although the exemption provision is not explicit in identifying trucks as the competition of interest, trucks are the only ubiquitous mode, and thus the ability to move a commodity by truck became the de facto standard for deciding whether a commodity should be considered inherently competitive and granted a categorical exemption.

automatically excluded from any further regulation during the life of the contract. This gave railroads the freedom to tailor their rates and service offerings on a shipper-by-shipper basis.¹⁵

Exempt traffic and contract shipments provide railroads with an opportunity to earn sufficient revenues for their long-term financial viability. Since the passage of the Staggers Rail Act, the share of shipments in the CWS designated “exempt” or “contract” has grown continuously, which is a major factor in explaining the improved financial condition of Class 1 railroads. This trend implies that the precision and overall quality of our estimated reasonable-price conditional distribution are likely to improve over time.

Nevertheless, shipments that are neither exempt nor under contract and therefore subject to the Staggers Rail Act provision to protect captive shippers from excessive shipment rates are likely to continue to exist, particularly for small shippers. Therefore, a mechanism for determining whether a shipment rate is excessive continues to be necessary. Our price benchmark provides a low-cost alternative to the current approach to addressing this statutory mandate, is ideally suited to ensuring that small shippers obtain relief from excessive prices, and does not require the use of an ad hoc allocated-cost methodology that implies irrational shipment pricing by railroads.

As shown in the Appendix, our estimate of $F(y | X)$ is straightforward to update when another year of CWS data becomes available. The values of y are deflated to real magnitudes to make them consistent with the existing values of y . As shown in the Appendix, computing the updated value of $\hat{F}(y^* | X^*)$ for a shipment with price y^* and shipment characteristics X^* simply requires computing an N -element sum, where N is the number of observations in the data set.

4. Choosing a Price Benchmark for a Shipment

Our use of a price benchmark is different from the typical use of the construct in a regulatory proceeding. Price cap regulation typically specifies a maximum price or set of maximum prices that a price-regulated firm is allowed to charge for all of its products. The prices are designed to allow the firm an opportunity to recover its total cost of production through prudent operation. Yardstick regulation determines these maximum prices by using information from a group of

¹⁵ The ability of a railroad to contract gave it substantial latitude to set rates differentially according to a shipper’s circumstances and willingness to pay. Railroads would be allowed not only to compete more aggressively for the newly exempted freight that is inherently competitive with trucks but also to set tariff rates for the nonexempt bulk commodities at levels equivalent to the most rail-dependent shipper’s willingness to pay. While shippers with more transportation options would be expected to refuse to pay the higher rate, a railroad could simply negotiate a discounted contract rate with terms tailored to its particular situation and willingness to pay. The price-differentiating railroad would then be able to set rates at levels that avoid pricing any traffic that makes a positive contribution to fixed cost recovery out of the market. As noted earlier, because of this incentive to extract rents but not price traffic out of the market, the short-run efficiency loss from railroads having pricing freedom is expected to be minimal. Indeed, limited deadweight-loss-associated railroad pricing is found by Grimm and Winston (2000, p. 65).

similar firms producing the same product. Again, the resulting yardstick price is used to set the maximum price that the firm can charge for its output.

Our application differs from these uses of a price benchmark because railroads sell thousands of different products, as distinguished by the product shipped and features of the origin and destination of the shipment, and a growing share of the shipments are provided at market-determined prices, whereas under price cap and yardstick regulation all of the firm's output is subject to maximum-price regulation. Our price benchmark determines only the level of an unreasonable price for shipments that are neither exempt nor under contract and therefore are still subject to the Staggers Rail Act regulatory mandate against excessive prices. Consistent with the logic of Demsetz (1968), our approach relies on market forces to determine whether a shipment price is the result of the actions of a dominant railroad and therefore unreasonable and worthy of further regulatory review to determine if the price is excessive and the shipper is worthy of rate relief.

Setting the value of the unreasonable price for a shipment involves balancing two risks. The first is the risk of incorrectly determining that the observed shipment price is unreasonable when it is reasonable, and the second is the risk of failing to determine that a truly unreasonable price is in fact unreasonable. Because our price benchmark is derived from the conditional distribution of reasonable prices given shipment characteristics, we can build on the theory of statistical hypothesis testing to determine the appropriate value of the price benchmark. Our price benchmark is analogous to a critical value for the test of the null hypothesis that a shipment price is reasonable versus the alternative that it is unreasonable.¹⁶

If the null hypothesis is rejected and the price is ultimately determined to be excessive, this raises the question of what the reset price should be. This decision also involves balancing two risks. The first is the risk of setting the price too low and increasing the probability that the railroad does not receive sufficient revenues to cover its production costs. The second is the risk that setting the price too high does not protect the shipper from an excessive price. The Staggers Rail Act anticipates the first risk by requiring the STB to make an annual determination of whether each Class 1 railroad is revenue adequate in the sense of earning sufficient revenues to recover its total cost of production.

The remainder of this section first presents the results of a simulation experiment to determine the value of the price benchmark that optimally balances—in the sense described below—the risks of failing to reject the hypothesis that a truly unreasonable price is reasonable versus the risk of falsely rejecting the null hypothesis for a price that is truly reasonable. Because we have no a priori reason to favor railroads over shippers or vice versa, we treat the error types symmetrically in determining the benchmark value for the reasonable price of a shipment.¹⁷ We

¹⁶ In this sense, our approach is consistent with the comparative institution approach of Demsetz (1969), because we do not claim any optimality properties for our approach, only that we believe it is better suited to solve the economic problem than existing approaches.

¹⁷ Similar to the process of setting the size of a statistical hypothesis test, our approach to setting the value of a reasonable shipment price can be modified to place greater relative weight on one of the two error types.

then use the results of this simulation to inform our choice of the value of the price benchmark for our assessment of the impact on annual railroad revenues and the extent of rate relief provided to the shipper from various choices for the reset price if the actual price is determined to be excessive using CWS data from four classes of products.

4.1. Simulation Experiment on the Selection of the Price Benchmark

To study the impact that the value of the price benchmark has on the probability of each type of misclassification error, we require an environment where we know with certainty whether a shipment price is competitively determined. The environment should have a realistic amount of variation in the price of a competitively provided product due to cost differences for supplying the product for the same realization of demand conditions. Conversely, the environment should also have a realistic amount of variation in the price of competitively provided products due to demand differences for the same realization of the cost of supplying the product.¹⁸

For the purposes of our simulation study, we assume that a number of hypothetical railroads supply $i = 1, \dots, N$ shipments, each with demand function $D_i(p) = Ap^{-\alpha_i}$ for $\alpha_i = \mathbf{z}_i'\delta + \eta_i$, where $\mathbf{z}_i, \delta \in \mathbb{R}^M$. Each element of the M -dimensional vector of demand shifters $\mathbf{z}_i \equiv (z_{i1}, z_{i2}, \dots, z_{iM})'$ is assumed to be an independent and identically uniformly distributed, $U(-1, 1)$, random variable. The M dimension for demand shifters is chosen to achieve a realistic level of variation in market prices for a given cost function realization for the railroads. We find that $M = 4$ is sufficient to achieve this goal. Let $\delta = (\frac{1}{4}, \dots, \frac{1}{4})' \in \mathbb{R}^M$ and $\eta_i \sim U(6, 7)$. The supports of the distributions of demand shifters \mathbf{z}_{ij} , where $j = 1, 2, \dots, M$, and values of the elements of δ are selected to yield market prices with reasonable markups over marginal cost for all price realizations.

For the railroads' cost function, let $C(q_1, q_2, \dots, q_N) = \sum_{i=1}^N [\mathbf{w}_i'\gamma + \varepsilon_i]q_i$, where \mathbf{w}_i is a K -dimensional vector of cost shifters and $\gamma \in \mathbb{R}^K$. We draw w_{ij} for $j = 1, 2, \dots, K$ as independent and identically distributed $U(-1, 1)$ random variables. Define $\mathbf{w}_i = (w_{i1}, w_{i2}, \dots, w_{iK})'$ and let $\varepsilon_i \sim U(4, 5)$. Set $\gamma = (\frac{1}{4}, \dots, \frac{1}{4})' \in \mathbb{R}^K$. We find that given the support of the distributions of the demand shifters and values of the elements of γ , setting $K = 5$ is sufficient to obtain realistic variation in market prices for a given demand function realization. This variation in the elements of \mathbf{w}_i is necessary to produce both high-price and low-price reasonable outcomes because of high and low values of marginal cost caused by variation in $\mathbf{w}_i'\gamma + \varepsilon_i$.¹⁹

¹⁸ We ran the simulation experiment described below for a variety of specifications of the dimensions and support of the demand and supply shifters and the marginal impacts of the demand and supply shifters on the elasticity of demand and marginal cost. The results support our conclusions about the appropriate choice of the price benchmark from percentiles of the conditional distribution of reasonable prices given shipment characteristics.

¹⁹ We experimented with different parameter values for the demand and cost functions and the dimensions of the demand and supply shifters. Our conclusions about the percentile of the conditional distribution of reasonable prices that minimizes the sum of squared misclassification errors are largely invariant to these modeling choices.

If we assume that railroads set each shipment price p_i to maximize the sum of profits over the N shipments,

$$\pi(p_1, p_2, \dots, p_N) = \sum_{i=1}^N D_i(p_i)p_i - C[D_1(p_1), D_2(p_2), \dots, D_N(p_N)]$$

yields prices for each of the $i = 1, 2, \dots, N$ shipments equal to

$$p_i = \left[\frac{-(z'_i \delta + \eta_i)}{-(z'_i \delta + \eta_i) + 1} \right] (w'_i \gamma + \varepsilon_i). \tag{1}$$

Equation (1) demonstrates that variations in the markups over marginal cost are driven by variation in the values of z_{ij} , and variations in marginal cost are driven by variation in w_{ij} . The combination of these observable sources of random variation along with the two unobservable sources of random variation in η_i and ε_i produces a realistic conditional distribution of shipment prices given z_{ij} and w_{ij} .

Translating the variables of this economic model into our notation for the conditional distribution of reasonable prices, we let $\mathbf{X}_i = (z'_i, w'_i)'$ equal the set of conditioning variables and $y_i = \ln(p_i)$ be the natural log of p_i .²⁰ We then use the N observations of y_i and \mathbf{X}_i to estimate the conditional distribution of reasonable prices given the shipment characteristics $\hat{F}(\ln(p) | \mathbf{X}, \mathbf{a})$, using the procedure summarized in Section 3 and presented in detail in the Appendix.

Our choice of the price benchmark is equivalent to selecting the value of the percentile of the conditional distribution of reasonable prices given shipment characteristics beyond which any observed price would be deemed unreasonable. Suppose that p^* is the price of a potentially unreasonably priced shipment with characteristics \mathbf{X}^* . If $1 - \alpha$, for $1 > \alpha > 0$, is the percentile of the distribution of $\hat{F}(p | \mathbf{X})$ beyond which prices are deemed to be unreasonable, then $PB(\alpha, \mathbf{X}^*)$ solves the equation $1 - \alpha = F(PB(\alpha, \mathbf{X}^*))$. If $p^* > PB(\alpha, \mathbf{X}^*)$, then the null hypothesis that the observed shipment price p^* is reasonable would be rejected. An equivalent decision rule is $\hat{F}(p^* | \mathbf{X}^*) > 1 - \alpha$, and then this null hypothesis would be rejected.

The simulation samples of truly reasonable and truly unreasonable prices used to determine the optimal value of α are constructed as follows. We repeat the process of drawing observations of $(z_i, w_i, \eta_i, \varepsilon_i)$ for $i = 1, 2, \dots, Q$ and compute p_i using equation (1) for all i . Then for every k value of p_i , we compute an unreasonable price \tilde{p}_i and replace p_i with \tilde{p}_i . The process used to compute these unreasonable prices is described below.

For each $i = 1, 2, \dots, Q$ in this test sample, define

$$I_i = \begin{cases} 1 & \text{if } p_i = \tilde{p}_i \\ 0 & \text{otherwise} \end{cases}.$$

²⁰ We estimate the cumulative distribution of the natural logarithm of price in recognition of the fact that prices are positive, and the distribution is positively skewed.

The indicator variable is equal to one if the i th price observation is truly unreasonable and zero if the observation is truly reasonable.

For each (p_i, z_i, w_i) combination in the test sample, we compute $\hat{F}(p_i | z_i, w_i)$ using the reasonable-price distribution estimated from the $N = 1,000$ price draws that solve equation (1) and are therefore reasonable. We then find the value of α that minimizes the sum of squared misclassification errors for our test sample of reasonable and unreasonable prices,

$$\sum_{i=1}^Q (I_i - \hat{I}_i)^2, \tag{2}$$

where \hat{I}_i is determined by the rule that \hat{I}_i equals one if $\hat{F}(p_i, z_i, w_i) > 1 - \alpha$ and zero otherwise. Depending on the value of α , the value of \hat{I}_i indicates whether the i th price exceeds the value of $PB(\alpha, X)$, the price benchmark for a shipment with characteristics X and that value of α . Ideally, we would like $\hat{I}_i = I_i$ both when I_i equals one and when it equals zero, which means that when the price is truly unreasonable it exceeds the price benchmark, and when it is truly reasonable it does not exceed the price benchmark. This would make our objective function equal 0 for all observations. Note that both types of misclassification errors, \hat{I}_i equals zero when I_i is one and \hat{I}_i equals one when I_i is zero, contribute the same value, 1, to the objective function. Solving for the value of α that minimizes expression (2) is the equivalent of finding the price benchmark function $PB(\alpha, X)$ that minimizes the sum of squared misclassification errors for observations $(y_i, X_i)'$ in our test sample of size Q of reasonable and unreasonable prices.

To compute unreasonable prices in our test sample, we change the distribution of η_i . We let

$$\tilde{p}_i = \left[\frac{-(z_i'\delta + \tilde{\eta}_i)}{-(z_i'\delta + \tilde{\eta}_i) + 1} \right] (w_i'\gamma + \varepsilon_i), \tag{3}$$

where $\tilde{\eta}_i \sim U(m, n)$. For all scenarios, $\eta_i \sim U(6, 7)$ for reasonable observations. The support of $\tilde{\eta}_i$ has the same range but is increasing closer to the support of η_i across the six scenarios listed in Table 1. The closer the support of $\tilde{\eta}_i$ is to the support of η_i , the more likely it is that our procedure will mistakenly classify unreasonable prices as reasonable, as shown in Table 1.

For all of the scenarios, we set $Q = 3,000$ and $k = 5$, which implies 600 excessive prices in each 3,000-observation test sample. Table 1 presents the value of α that minimizes the sum of squared misclassification errors for each distribution of $\tilde{\eta}_i$ and the type I errors (reasonable prices classified as unreasonable observations), type II errors (unreasonable price observations classified as reasonable), and percentages of misclassified observations (the sum of type I and II errors).

As the support of the distribution of $\tilde{\eta}_i$ approaches the support of the distribution η_i , the distribution of excessive prices is closer to the distribution of reasonable prices. The percentage of observations in our test sample that are misclassified also rises. However, even when the supports of $\tilde{\eta}_i$ and η_i are virtually the

Table 1
Simulation Results: Choice of the Optimal Value of α

$\tilde{\eta}_i$	α	Type I Error	Type II Error	Percentage Misclassified
$U(3.75, 4.75)$.0478	34	339	12.4
$U(4, 5)$.0627	58	365	14.1
$U(4.5, 5.5)$.0478	34	481	17.2
$U(5, 6)$.0303	17	553	19.0
$U(5.5, 6.5)$.0133	2	586	19.6
$U(5.75, 6.75)$.0133	2	590	19.7

same, (5.75, 6.75) versus (6, 7), less than 20 percent of the observations in the test sample are misclassified. Finally, for all of the scenarios considered, the value of α that minimizes the sum of squared misclassification errors lies in the interval (.0133, .0627).

These results demonstrate that the greater the overlap between the support of the distribution of reasonable prices and the support of the distribution of unreasonable prices, the smaller is the value of α that minimizes the sum of the squared misclassification errors. The results suggest that the value of the optimal α is unlikely to be larger than .06 and smaller than .01.

Figures 1 and 2 present graphs of the values of $\hat{f}(\ln(p) | \mathbf{X})$ and $\hat{F}(\ln(p) | \mathbf{X})$ for the pairs of price and shipment characteristics in our test samples for the scenarios with the smallest and largest overlaps in the supports of the distributions of truly reasonable and truly unreasonable prices, the values of $\tilde{\eta}_i$ in the first and last rows of Table 1.²¹ Figures 1 and 2 and Table 1 show that the major cost in terms of misclassification errors as the supports of the distributions of reasonable and unreasonable prices come closer is a substantial increase in type II errors—failing to find that a truly unreasonable price is unreasonable. For the scenarios considered, the frequency of type I errors—concluding that a reasonable price is unreasonable—remains very low.

5. Determining a Reasonable Price for a Shipment

Using the above results, we now implement our price benchmark procedure using data from the CWS to determine the appropriate reasonable price to use if an actual price is found to be excessive by the STB. To this end, we estimate separate conditional distributions for four broad commodity groups: petroleum products, farm products, coal, and chemical products. We apply our price benchmark approach to all potential unreasonable prices for $\alpha = .05$ and $\alpha = .01$. We then consider the total revenue implications of resetting unreasonable prices to various reasonable levels for railroads and compute the average amount of regulatory relief our procedure provides to shippers.

These reasonable-price conditional distributions can be estimated for more

²¹ Graphs of $\hat{f}(\ln(p) | \mathbf{X})$ and $\hat{F}(\ln(p) | \mathbf{X})$ for the prices for the remaining scenarios are in the Online Appendix.

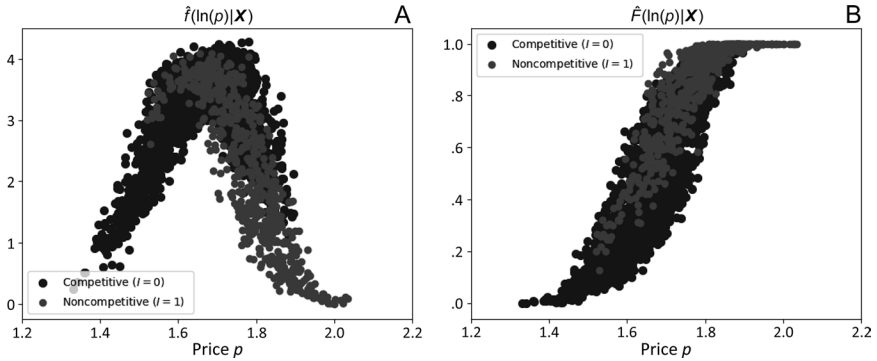


Figure 1. Conditional distributions with $\eta_i \sim U(6, 7)$ and $\tilde{\eta}_i \sim U(3.75, 4.75)$. A, Probability distribution function; B, cumulative distribution function.

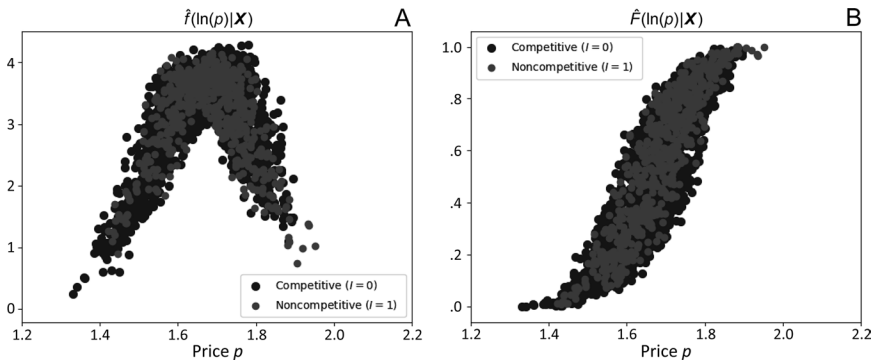


Figure 2. Conditional distributions with $\eta_i \sim U(6, 7)$ and $\tilde{\eta}_i \sim U(5.75, 6.75)$. A, Probability distribution function; B, cumulative distribution function.

commodities and for narrower product groups (for example, grain or hazardous materials) as long as there are sufficient observations to obtain credible estimates of $\hat{F}(y | \mathbf{X})$. Once the conditional distribution is estimated for each commodity, the STB can use it and the procedure described above to determine whether a shipper is being charged an unreasonable price for a movement or set of movements given the shipment characteristics and a value of α . As we discuss below, the results from our simulation study imply that the value of α can be different for different commodities depending on the location of the support of the distribution of reasonable prices versus the location of the support of the distribution of unreasonable prices, but all values of α should be in the range (.01, .06).

A number of commenters on a previous draft raised the point that railroads would likely change their pricing strategy for shipments that go into the computation of the conditional distribution of reasonable prices given shipment char-

acteristics. However, this response to the implementation of our price benchmark approach implies that railroads are not currently maximizing the profits they earn from pricing shipments. A railroad changing its pricing behavior in response to the implementation of our price benchmark mechanism could maximize the discounted present value of its expected profits if giving up some profits in the current period achieves a large enough increase in expected profits in future periods. However, we believe that our approach to choosing the price benchmark makes railroads unlikely to change their behavior. As shown in Section 4, the price benchmark is an extreme percentile of the conditional distribution of reasonable prices given shipment characteristics. Even if railroads increase prices on the vast majority of their reasonably priced shipments, this behavior is unlikely to impact significantly the upper α th percentile of the conditional distribution for values of α in the range of .01 to .06 recommended by our simulation study. This logic implies that the expected future payoff to the railroad is unlikely to compensate for lost profits in the current period caused by the railroad setting a higher shipment price to increase the price benchmark used to determine if a price is unreasonable in the future.

Estimating our conditional reasonable-price distribution relies primarily on data from the CWS. The dependent variable y_i is the natural logarithm of the average revenue per ton-mile for the shipment deflated by the gross domestic product price deflator.²² This variable is the revenue received from a shipment divided by the product of the number of tons in the shipment and the distance traveled. Revenues are the sum of freight revenues (transportation-related revenues), miscellaneous charges, and fuel surcharges.²³ In the calculation for ton-miles, billed weight is used for tons, and distance is calculated as the total miles traveled for the shipment.

The elements of \mathbf{X} , the vector of shipment characteristics, are shipment distance (X_1), shipment size in number of cars (X_2), the number of railroads involved in the movement (X_3), the number of class 1 railroads within 10 miles of the origin (X_4), the number of class 1 railroads within 10 miles of the destination (X_5), a dummy indicating whether the shipper owns the rail cars used for the shipment (X_6), a dummy indicating that there is no water port within 50 miles of the origin (X_7), and a dummy indicating that there is no water port within 50 miles of the destination (X_8). As noted in the Appendix, it is straightforward to add variables to the vector of shipment characteristics \mathbf{X} .

The elements of \mathbf{X} in this implementation are selected on the basis of two factors: previous empirical research on the determinants of shipment rates and the availability of the variables in the CWS and other public data sets (see Boyer 1987;

²² The logarithmic function transforms a positively skewed distribution of prices into a more symmetric distribution.

²³ Fuel surcharges were introduced by railroads in 2003 but are reported in different Carload Waybill Sample (CWS) fields by different railroads. Some railroads include them in the freight revenue field, and others include them in the miscellaneous revenue field. From 2009 forward, the CWS has had a separate field for fuel surcharges. Therefore, our solution is to use total revenues, including fuel surcharges, for the shipment as our shipment revenue variable.

Barnekov and Kleit 1990; McFarland 1989; Burton 1993; Wilson 1994; Dennis 2001; Schmidt 2001; MacDonald 1987, 1989; Grimm, Winston, and Evans 1992; Burton and Wilson 2006). The continuous variables—distance, size, and number of railroads—are measured in natural logarithms to make their marginal distributions more symmetric. (In preliminary versions of this analysis, we experimented with different distances for constructing X_4 , X_5 , X_7 , and X_8 and obtained quantitatively similar empirical results.) Finally, fixed effects are included for the year of the movement, the primary railroad in the movement, and the five-digit Standard Transportation Commodity Code categories. Each shipment in the CWS has an expansion factor EF_i , which gives the STB's estimate of the number of shipments in the population of annual shipments with the same observable characteristics. Results from estimating this reasonable-price conditional distribution function as described in the Appendix for petroleum products, farm products, coal, and chemical products data for 2000–2013 from the CWS are presented below.

5.1. Unreasonable-Price Analysis

For each product group, the sample of reasonable-price observations is first used to compute an estimate of $\hat{F}(e^p | es^{X^c})$ defined in the Appendix. Then the shipment price for each observation in the potentially unreasonable-price sample for each product group is compared with the benchmark price computed as described above for $\alpha = .05$ and $\alpha = .01$. Table 2 presents the estimated population percentages of shipments for each product group that are found to be unreasonable for each value of α . The percentages of unreasonable prices are the ratios of the sum of the expansion factors for shipments that exceed the price benchmark for α divided by the total number of shipments. Because the expansion factor gives the estimated number of shipments in the population that each shipment in the CWS sample represents, each number is an estimate of the population percentage of shipments that violate that price benchmark for that value of α .

There are likely sufficient observations in the reasonable-price sample for each product group to obtain a precise estimate of $F(y | \mathbf{X})$. There are 50,340 reasonable-price observations for petroleum products, 53,205 for farm products, 285,976 for coal, and 356,187 for chemical products. Across all years and all products except coal, the frequency of unreasonable shipment prices for $\alpha = .05$ is less than 5 percent except in 2006 for petroleum products (5.04 percent) and 2002 for chemical products (6.81 percent). For $\alpha = .01$, for these three product groups the frequency of unreasonable prices is less than 1.5 percent, and for most years it is less than 1 percent. The above results for the frequency of unreasonable shipments and the greater heterogeneity of the goods shipped in these three product groups relative to coal make a value of $\alpha = .05$ more appropriate for these three product groups.

For coal the annual frequency of unreasonable-price observations for $\alpha = .05$ is as high as 26 percent in 2006. With $\alpha = .01$ the annual frequency of excessive observations never exceeds 9 percent. A number of trends in the US coal sec-

Table 2
 Estimated Populations of Unreasonable-Price Percentages

Year	Petroleum Products			Farm Products			Coal			Chemical Products		
	Shipments	$\alpha = .05$	$\alpha = .01$	Shipments	$\alpha = .05$	$\alpha = .01$	Shipments	$\alpha = .05$	$\alpha = .01$	Shipments	$\alpha = .05$	$\alpha = .01$
2000	150,276	2.26	.27	107,502	1.55	.69	313,470	10.53	8.14	948,045	3.81	.42
2001	120,566	3.28	.91	101,372	2.68	1.07	206,174	3.97	1.48	750,680	4.47	.65
2002	84,216	4.81	1.49	100,725	2.30	1.04	211,770	10.65	6.07	572,671	6.81	1.17
2003	82,932	3.44	.90	90,607	2.16	.72	162,456	3.82	1.71	542,324	3.68	.77
2004	80,452	3.88	.45	85,868	2.51	.68	153,933	3.72	1.71	605,448	2.62	.36
2005	94,296	3.13	1.03	95,520	2.46	1.17	109,765	7.14	3.11	635,763	3.60	.88
2006	108,212	5.04	1.17	98,426	1.97	.61	35,110	26.25	8.95	664,846	3.09	.47
2007	102,844	4.94	1.13	94,673	1.87	.34	33,924	23.26	8.49	501,475	3.96	.48
2008	94,992	3.93	1.02	87,670	2.74	1.32	33,024	16.95	5.36	451,544	2.71	.60
2009	88,812	4.17	1.12	82,393	2.50	1.47	24,077	16.96	6.29	407,162	3.26	.69
2010	81,804	4.73	1.35	81,708	1.83	1.05	20,374	21.66	7.61	413,085	3.26	.74
2011	87,312	3.81	.96	78,675	2.04	1.11	18,557	22.98	6.67	395,827	3.19	.84
2012	98,588	2.40	.74	74,550	2.22	1.24	12,147	23.00	7.87	305,394	4.24	1.08
2013	101,256	3.07	.64	66,206	2.67	1.19	11,652	14.45	6.92	282,837	3.90	.91
All	1,376,558	3.70	.91	1,245,895	2.24	.96	1,346,433	9.17	4.76	7,477,101	3.79	.68

tor caused the substantial reduction in the estimated annual population of shipments. First, an increasing share of coal deliveries came from surface mines west of the Mississippi River. These shipments involved a substantially larger number of coal cars: unit trains that move western coal can involve hundreds of coal cars, whereas earlier movements involving eastern coal from underground mines often involved fewer than 10 coal cars. These trends and the fact that coal shipments are more homogeneous than shipments of the other three product groups imply significant overlap between the support of the distribution of reasonable prices and the support of the distribution of unreasonable prices. From the results of our simulation study in Section 4.1, this result argues in favor of a value of α in the neighborhood of .01 for coal.

5.2. Average Revenue Change from Reasonable Pricing

In Table 3, we assess the impact of resetting shipment prices that are deemed to be unreasonable to different features of the conditional reasonable-price distribution on the revenues that railroads earn from shipments in the four product groups over our sample period. We consider three possible reasonable prices for shipments deemed unreasonable on the basis of our price benchmark for a given value of α . First, we reset an unreasonable price to the conditional mean of the reasonable-price distribution. Second, we set it equal to the conditional median of the reasonable-price distribution. Finally, we set it equal to $PB(\alpha, \mathbf{X})$, the price benchmark for that shipment. We show the percentage change in total revenues for the four product groups over our sample period for $\alpha = .1, .05, .01$.

Particularly for the reasonable price set equal to our price benchmark, the aggregate revenue implications for railroads of resetting the actual price to this price are less than 1.2 percent for α less than or equal to .05 for petroleum products, farm products, and chemical products. Even for reasonable prices equal to the conditional mean and median, the revenue reductions are less than 3.3 percent for α less than or equal to .05.

For coal, resetting unreasonable prices to any of the three features of the reasonable-price distribution for the case of $\alpha = .05$ implies at most a 9.3 percent reduction in annual revenues. For the case of $\alpha = .01$, the largest percentage reduction in annual revenues is 3.4 percent, which provides further evidence that $\alpha = .01$ is likely to be the appropriate choice for the coal sector.

5.3. Average Price Change from Reasonable Pricing

To assess the extent to which shippers obtain rate relief by resetting an unreasonable price to one of the three reasonable prices, we compute the average difference between the actual unreasonable price and the reasonable price for each measure of a reasonable price. Table 4 reports the average value of the price differences. The last line for each product group reports the average value of actual prices for all of the prices in the test sample deemed to be unreasonable.

These results demonstrate that even for the case of the reasonable price equal

Table 3
 Percentage Changes in Railroad Revenues by Reasonable Price

	Petroleum Products			Farm Products			Coal			Chemical Products		
	$\alpha = .10$	$\alpha = .05$	$\alpha = .01$	$\alpha = .10$	$\alpha = .05$	$\alpha = .01$	$\alpha = .10$	$\alpha = .05$	$\alpha = .01$	$\alpha = .10$	$\alpha = .05$	$\alpha = .01$
Mean p	-3.19	-2.20	-.82	-.73	-.83	-.66	-4.42	-3.50	-2.16	-.81	-1.53	-.63
Median p	-5.61	-3.07	-.97	-1.88	-1.23	-.73	-13.28	-9.31	-3.36	-5.62	-3.22	-.81
Threshold p	-2.24	-1.13	-.32	-1.20	-.80	-.37	-6.09	-3.60	-.97	-2.37	-1.19	-.29

Table 4
Average Shipment Price Changes for Reasonable Prices

	Petroleum Products			Farm Products			Coal			Chemical Products		
	$\alpha = .10$	$\alpha = .05$	$\alpha = .01$	$\alpha = .10$	$\alpha = .05$	$\alpha = .01$	$\alpha = .10$	$\alpha = .05$	$\alpha = .01$	$\alpha = .10$	$\alpha = .05$	$\alpha = .01$
Mean p	-10.03	-14.87	-24.97	-.61	-1.02	-1.87	-.12	-.12	-.19	-5.13	-8.27	-16.61
Median p	-11.62	-16.44	-26.58	-.72	-1.13	-2.00	-.16	-.16	-.25	-6.64	-9.81	-18.13
Threshold p	-4.01	-5.25	-7.51	-.30	-.53	-.89	-.05	-.05	-.08	-2.40	-3.31	-5.80
Average p of unreasonable prices	13.59	18.38	28.56	1.06	1.50	2.43	.25	.25	.37	8.55	11.75	20.04

to our price benchmark, the average price change is a significant fraction of the average price deemed to be unreasonable. For example, for petroleum products and $\alpha = .05$, the average unreasonable price is \$18.38, and the average difference between this price and the reasonable price using our price benchmark is \$5.25, which implies an average price reduction of more than 28 percent. For farm products and $\alpha = .05$, the average unreasonable price is \$1.50, and the average price difference is \$.53, which implies an average price reduction of more than 35 percent.

For coal, the average price change is 5 cents for $\alpha = .05$. The average unreasonable price for this scenario is 25 cents, an average price reduction of 20 percent, a significantly lower percentage change than for petroleum products or farm products. Nevertheless, given the large shipment volumes involved, a 20 percent reduction in price implies significant rate relief. For chemical products, the average price change is \$3.31 for $\alpha = .05$. The average unreasonable price for this value of α is \$11.75, an average price reduction of 28 percent, which is in the neighborhood of the values obtained for petroleum products and farm products and significantly larger than the value for coal.

5.4. Analysis Restricted to Convex Hull of Reasonable Price Observations

To address concerns that the estimates of $\hat{F}(y | \mathbf{X})$ might be imprecise because the values of \mathbf{X} for shipments in the potentially unreasonable-price sample are unrepresentative of the values of \mathbf{X} in the reasonable-price sample, we perform the analysis in Tables 2, 3, and 4 but restrict the set of shipments in the potentially unreasonable-price sample to those with vectors of shipment characteristics \mathbf{X} that are contained in the convex hull of the vectors of shipment characteristics in the reasonable-price sample.²⁴ The results are shown in Tables OA1–OA3 in the Online Appendix.

The vast majority of expansion-factor-weighted shipments in our full unreasonable-price sample have a vector of shipment characteristics in that convex hull. For petroleum products, farm products, and chemical products, the percentages of unreasonable prices for both values of α tend to be very similar to the corresponding values for the full sample for almost all of the years in our sample. For coal, restricting the sample to shipments with characteristics in that convex hull results in a significant fraction of shipments being excluded from the unreasonable-price analysis.²⁵ Nevertheless, the percentage of shipments found to be unreasonable in the restricted sample is very similar to the percentages found to be unreasonable in the full sample for most years.

Table OA2 shows that for petroleum products, farm products, and chemical products, the percentage revenue changes for the restricted sample is very simi-

²⁴ The convex hull of the set \mathbf{X} of $i = 1, 2, \dots, N$ vectors $\mathbf{X}_i = \{x_{i1}, x_{i2}, \dots, x_{iM}\}$ of dimension M is

$$\text{Conv}(\mathbf{X}) = \{Z | Z = \sum_{i=1}^N \alpha_i \mathbf{X}_i, \alpha_i \geq 0 \text{ for } i = 1, 2, \dots, N, \sum_{i=1}^N \alpha_i = 1\}.$$

²⁵ Recall the earlier discussion of the massive shift during the mid-2000s from rail shipments with few cars of eastern coal to rail shipments with many cars of western coal.

lar to those for the full sample in Table 3. For coal, the changes for the restricted sample are uniformly smaller than those for the full sample for all values of α and choices of the mitigated shipment prices.

Table OA3 shows that for the restricted sample of all product groups, the absolute values of the average price change are typically smaller than those for the full sample. However, for all product groups the mean unreasonable price is slightly smaller for the restricted sample, and so the average price changes relative to this mean are similar to those for the full sample.

Taken together, these results argue against the concern that the support of the distribution of the vector of shipment characteristics of the potentially unreasonable-price sample is significantly different from those of the reasonable-price sample for petroleum products, farm products, and chemical products. The case of coal suggests modifying the unreasonable-price determination procedure to require potentially unreasonable prices to lie in the convex hull of the vectors of shipment characteristics from the reasonable price sample. This logic provides further support for using $\alpha = .01$ to compute the price benchmark for coal to account for the fact that the distributions of the vectors of shipment characteristics do not completely overlap. Substantially greater evidence against the hypothesis that the observed price is reasonable is necessary to make an unreasonable-price determination.

A number of conclusions emerge from our simulation study and the application of our benchmark price to data from the CWS. First, our simulation study finds that values of α between .01 and .06 appear to minimize the sum of squares of misclassification errors for the types of conditional distributions of reasonable prices and of unreasonable prices likely to be encountered in practice. Second, for those values of α , the vast majority of shipment prices in our simulation test sample are correctly classified as reasonable when they are truly reasonable.²⁶ Third, for those values of α , even resetting the unreasonable price to the conditional mean or conditional median of the reasonable-price distribution is likely to have a small adverse impact on the revenues earned by carriers of the four product categories. Fourth, resetting the value of an unreasonable price to our price benchmark has the smallest adverse impact on railroad revenues. Fifth, for all products we find that resetting unreasonable prices to any of our three features of the reasonable-price distribution produces economically meaningful price reductions for the affected shippers. For $\alpha = .05$, the smallest average percentage price reduction for mitigated shipments relative to the average unreasonable price across the four product groups is 20 percent. The results presented in this section suggest that our price benchmark approach can be a low-cost administrative approach for the STB to carry out its statutory mandate to protect shippers from excessive prices, while at the same time not adversely impacting the ability of the railroads to achieve the aggregate revenues necessary for their long-term financial viability.

²⁶ When the supports of the distribution of unreasonable prices and reasonable prices overlap, most of the misclassification errors are due to classifying unreasonable prices as reasonable prices.

6. Use of a Price Benchmark Mechanism in the Regulatory Process

There are a variety of ways to use our reasonable-price conditional distribution function to carry out the STB's mandate to protect captive shippers from excessive prices. We believe that a benchmark price approach should replace the $R/VC \leq 180$ test in the decision to declare a price the result of market dominance, the first step in the STB rate relief process. The fact that our approach compares the price charged to an extreme percentile of the reasonable-price distribution is consistent with the economic logic that a railroad is market dominant if it charges an unreasonable price for a shipment. Other factors could be considered by the STB in making the determination of market dominance, but it seems imprudent to use $R/VC \leq 180$ given that the VC measure is based on an ad hoc allocated-cost measure that is sensitive to what fixed and common costs are allocated to a shipment.

We recommend a final-offer-arbitration process for the second stage of determining whether an unreasonable price determined to be the result of market dominance is excessive and worthy of regulatory relief.²⁷ The reasonable-price distribution could be used in two ways in this process: to provide evidence for or against the shipment price being excessive and to provide evidence for or against the value of the price to which the actual price should be reset when a finding of an excessive price is made.

Using the price benchmark approach to replace the $R/VC \leq 180$ test for an excessive price has the advantage of ending the STB's reliance on URCS VC measures in finding an excessive price.²⁸ The Online Appendix compares the $R/VC \leq 180$ test with our price benchmark test. For each shipment in our potentially unreasonable-price sample, an indicator variable is set equal to one if this test is violated and zero otherwise. For a given value of α , each shipment is placed into one of four categories: both the $R/VC \leq 180$ test and the benchmark price test are violated, the $R/VC \leq 180$ test is not violated and the benchmark price test is violated, the $R/VC \leq 180$ test is violated and the benchmark price test is not violated, and neither test is violated.

For the four product groups and $\alpha = .05$ and $\alpha = .01$, the majority of violations of our price benchmark test also violate the $R/VC \leq 180$ test. However, there are also instances in which the benchmark price test is violated but the $R/VC \leq 180$ test is not. For petroleum products and chemical products, less than 1 percent of the shipments violated our price benchmark test but did not violate the $R/VC \leq 180$ test. For farm products the percentages were somewhat higher but never larger than 1.3 percent. For coal the percentages averaged less than 4 percent across all years. These results imply that there are instances in which our approach yields a finding of dominance and the current approach does not.

However, the major difference between the two approaches is the extremely

²⁷ This is consistent with the recommendations for reform of the rate relief process in TRB-NAS (2015).

²⁸ This logic also calls into the question the usefulness of the Three-Benchmark approach, which relies on the Uniform Railroad Costing System VC of a shipment to make an excessive-price determination, a topic we leave to future research.

high frequency that the $R/VC \leq 180$ test is violated and our price benchmark is not violated. For petroleum products and chemical products, approximately 45 percent of potentially unreasonable-price shipments violate the $R/VC \leq 180$ test but do not violate the price benchmark test. The percentages are above 18 percent for farm products and above 14 percent for coal. These results are consistent with the logic that the $R/VC \leq 180$ test is significantly less discerning in distinguishing unreasonable prices from reasonable prices than our price benchmark approach. This implies that the STB must devote significantly more time and effort to examining factors other than the results of the $R/VC \leq 180$ test in making a determination of market dominance than would be the case for our price benchmark approach.

Changing the current excessive-price determination process to rely on our price benchmark approach would likely require legislation, which raises the question of who should set the value of α that determines the value of the price benchmark. Similar to the case of the $R/VC \leq 180$ test, the value of α could be set in the legislation that implements the price benchmark. Alternatively, the law could provide legislative guidance to the STB in setting the value of α . For example, the law could direct the STB to set α to minimize an estimate of the sum of squared misclassification errors.

Even without a legislated change in the excessive-price determination process, our price benchmark approach can provide valuable input to both steps of the current rate relief process. Given the large number of violations of the $R/VC \leq 180$ test shown in the Online Appendix, the price benchmark can provide more discerning input to determining market dominance. The empirical results in Section 4 suggest that the conditional reasonable-price distribution can inform both the excessive-price decision-making process and the process of setting the value of the mitigated price if regulatory relief is granted.

One factor that the STB may wish to consider before concluding that a violation of our price benchmark implies that the price charged is the result of actions by a market-dominant railroad is that an important factor is left out of vector X of product and shipment characteristics in computing the conditional distribution of reasonable prices. If the STB concludes that there is sufficient evidence that certain product or shipment characteristics are not included in X and that these characteristics are likely to explain the high price, then this violation of our price benchmark should not result in a finding of market dominance. This logic could also be applied in the final-offer-arbitration process. The arbitrator could determine if the factors proposed by the railroad explain the higher price charged by the railroad and if therefore the shipper is not entitled to rate relief.

This arbitration process also could provide input to the computation of future price benchmarks. If a factor not included in the vector of observed shipment characteristics is found by the arbitrator to explain the higher price, the STB could require data on that factor to be compiled for all future shipments sampled for the CWS data. That factor could then be incorporated into the vector of observed characteristics used to compute the conditional distribution of reasonable prices. For example, a number of railroads argue that hazardous materials

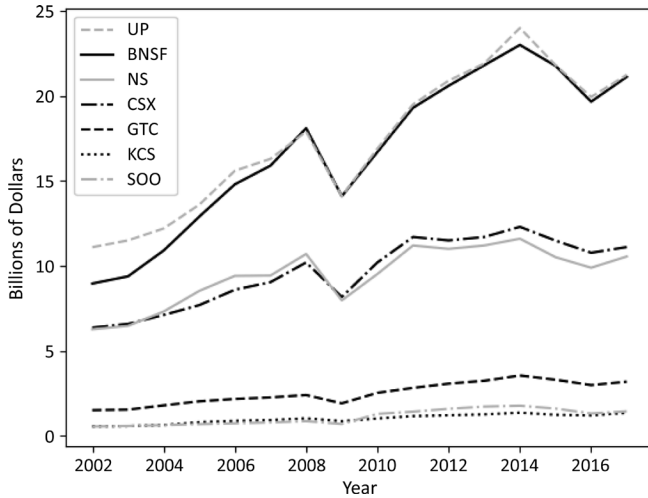


Figure 3. Annual operating revenues for US Class 1 railroads, 2002–2017

are more expensive to move and therefore charge higher prices to ship them. On the basis of arbitration results for this issue, the STB could require shippers to report the dimensions of hazardous materials in a shipment in CWS data. These observable factors could be incorporated into the vector of product and shipment characteristics used to compute the conditional distribution of reasonable prices.

A final issue that the price benchmark approach can address is the impact that the selection of the feature of the reasonable price distribution has on the annual revenue adequacy of the railroad. As shown in Section 5, resetting an unreasonable price using a percentile of the conditional distribution of reasonable prices allows an analysis of the annual revenue implications of different choices of α for determining the reasonable price. Smaller values of α imply a larger value of the reasonable price and therefore a smaller reduction in annual revenues from resetting excessive prices.

This logic suggests another factor to consider in setting the value of α : the year-to-year volatility in rail revenues from movements involving the product under consideration. Figure 3 plots the annual operating revenues for the seven Class 1 railroads in the United States from 2002 to 2017.²⁹ The year-to-year variation in revenues excluding the financial crisis in 2008–9 provides guidance for selecting the value of α . According to these graphs, an annual revenue change of 5 percent is consistent with the year-to-year variation from trend growth in revenues over time for all of the Class 1 railroads.

This logic implies that product categories that typically experience less year-to-

²⁹ In Figure 3, BNSF is Burlington Northern and Santa Fe, CSX is CSX Transportation, GTC is Grand Trunk Corporation, KCS is Kansas City Southern Railway, NS is Norfolk Southern Corporation, SOO is SOO Line Corporation, and UP is Union Pacific.

year variation in revenues relative to the trend should have lower values of α than products that experience more year-to-year variation. Among our four product categories, we would expect coal to typically experience the least year-to-year variation in revenues, given that historically coal has been used to produce base-load electricity. Petroleum, farm products, and chemicals are likely to have higher year-to-year variation in product-level revenues. This logic implies that coal should have a smaller value of α than the other three products. This reinforces our earlier recommendation of a smaller value of α based on our finding that the frequency that the vector of shipment characteristics of potentially unreasonable-price shipments lies in the convex hull of those for reasonable-price shipments is significantly lower for coal relative to the other three product classes.

7. Concluding Comments

The fact that a growing share of rail shipments are moving at rates determined under effectively competitive conditions according to the STB presents an opportunity to use the data to construct a conditional distribution of reasonable prices given shipment characteristics that can be used to determine whether shipment rates are unreasonable and therefore the result of the actions of a market-dominant railroad. The computation of the price benchmark can be automatically updated each year given a sample of shipment prices and observable characteristics and their expansion factors from the annual CWS. Moreover, the conditional distribution of reasonable prices can be updated to condition on additional observable characteristics that are found to explain shipment prices. Even without a legislative change in the current approach to providing rate relief to captive shippers, our price benchmark approach can provide useful input to this process in a manner that provides significant rate relief to shippers without adversely impacting the revenue adequacy of the railroad subject to regulatory intervention.

Appendix

Conditional Distribution Estimation Procedure

This Appendix presents our kernel regression estimate of $F(y | \mathbf{X})$. Kernel regression is a nonparametric method for estimating the conditional mean function of one element of a random vector, in this case $I(Y_i \leq y)$, the indicator function that equals one if random variable Y_i is less than a real number y and zero otherwise, given the remaining elements of the random vector \mathbf{X} . Note that the conditional expectation of $I(Y_i \leq y)$ given \mathbf{X}_i , $E[I(Y_i \leq y) | \mathbf{X}_i]$, is equal to $F(y | \mathbf{X}_i)$.

In general, the kernel or local smoothing estimate of the conditional expectation of a random variable Y given a vector \mathbf{X} , $E(Y | \mathbf{X})$ is a weighted average of all of the values of Y_i in the sample, where larger weights are associated with Y_i values that have \mathbf{X}_i values that are closer to \mathbf{X} , the point of evaluation of $E(Y | \mathbf{X})$. This logic implies that if there were a finite number of possible values of \mathbf{X} , $\hat{E}(Y | \mathbf{X}^*)$ would simply equal the sample mean of the Y_i values for all observations with

a value of X equal to X^* . If some of the elements of X are continuously distributed random variables, a local smoothing approach must be employed to obtain a consistent estimate of $E(Y | X)$ for all possible values of X . This approach yields a consistent estimate of $E(Y | X)$ without imposing any functional form restrictions on the relationship between Y and X . Under the regularity conditions given in Bashtannyk and Hyndman (2001) or Hall, Wolff, and Yao (1999), the local smoothing or kernel regression estimator is a consistent estimate of the population conditional distribution Y given X .

To implement our nonparametric conditional distribution estimator, we first divide the vector X into two groups of variables: continuous (X^c) and binary (X^d), which include the categorical variables used for fixed effects. Using the data set composed of N observations classified as being reasonably priced (in this case exempt and contract shipments), we regress y_i on X_i^d and compute the residuals e_i^p and regress each element of X_i^c on X_i^d and compute the vector of residuals $e_i^{X^c}$ for observations $i = 1, 2, \dots, N$. Because the CWS is a stratified sample, we run each regression using weighted least squares with the weight for observation i equal to the expansion factor for that observation. Let EF_i equal the expansion factor associated with the i th observation, which is equal to the number of shipments in the population of annual shipments that have the same observable characteristics X as this shipment. All of the observations of each element of $e_i^{X^c}$ are scaled by the sample standard deviation of that element of e^{X^c} to compute $es_i^{X^c} = S^{-1}e_i^{X^c}$, where S is the diagonal matrix of the sample standard deviations of the elements of $e_i^{X^c}$. Because the simulation study in Section 4 has only continuous shipment characteristics, this step in the estimation process is unnecessary.

We then estimate the conditional distribution $F(e^p | es^{X^c})$ using the Nadaraya (1964) and Watson (1964) kernel regression estimator that accounts for stratified sampling of shipments in the CWS data:

$$\hat{F}(e^p | es^{X^c}, a) = \frac{\sum_{i=1}^N (EF_i)K_a(es^{X^c} - es_i^{X^c})I(e_i^p \leq e^p)}{\sum_{i=1}^N (EF_i)K_a(es^{X^c} - es_i^{X^c})}$$

where $I(e_i^p \leq e^p)$ is one if e_i^p is less than or equal to e^p and zero otherwise. Definitions of the function $K_a(es^{X^c} - es_i^{X^c})$ and the vector a and how the value of a is chosen are discussed below.³⁰

To construct $K_a(\cdot)$ we use the Epanechnikov kernel:

$$K(x) = \begin{cases} \frac{3}{4}(1 - x^2) & \text{for } |x| < 1 \\ 0 & \text{otherwise} \end{cases}$$

³⁰ For the simulation study reported in Section 4, we compute

$$\hat{F}(y | X, a) = \frac{\sum_{i=1}^N K_a(X - X_i)I(Y \leq y)}{\sum_{i=1}^N K_a(X - X_i)}$$

and use leave-one-out cross validation to compute a .

which yields

$$K_a(\mathbf{es}^{X^c} - \mathbf{es}_i^{X^c}) = \prod_{j=1}^J \frac{1}{a_j} K\left(\frac{es_j^{X^c} - es_{ij}^{X^c}}{a_j}\right),$$

where $\mathbf{a} = (a_1, a_2, \dots, a_j)$, $es_j^{X^c}$ is the j th element of \mathbf{es}^{X^c} and $es_{ij}^{X^c}$ is the j th element of the i th observation of $\mathbf{es}_i^{X^c}$. Other choices of the kernel function $K(t)$ produce similar estimates of $F(e^p | \mathbf{es}^{X^c})$. Once the vector \mathbf{a} is selected, our estimate of $F(e^p | \mathbf{es}^{X^c})$ can be computed given a size N sample of $(e_i^p, \mathbf{es}_i^{X^c})'$ and associated expansion factors $EF_i, i = 1, 2, \dots, N$.

We choose values of the elements of \mathbf{a} according to the bootstrap bandwidth selection approach analyzed by Bashtannyk and Hyndman (2001) and originally recommended by Hall, Wolff, and Yao (1999).³¹ Our procedure for computing the bandwidth parameter vector \mathbf{a} for our kernel regression estimator of $F(e^p | \mathbf{es}^{X^c})$ first fits a rich polynomial regression to predict e_i^p using polynomials in the elements of $es_i^{X^c}$:

$$e_i^p = \beta_0 + \sum_{j=1}^J \beta_{j1} es_{ij}^{X^c} + \dots + \beta_{jk} (es_{ij}^{X^c})^k + \sigma \varepsilon_i,$$

where ε_i is a group of regression errors that are assumed to be independently and identically distributed $N(0, 1)$ random variables, and k , the order of polynomial, is determined by the information criterion in Akaike (1973). Given the estimated values of β and σ , we then form a parametric estimator $\tilde{F}(e^p | \mathbf{es}^{X^c})$ from the model under the assumption that the values of ε_i are independent and identical normally distributed random variables. Then we simulate $l = 1, 2, \dots, L$ bootstrap data sets of size N , $\mathbf{e}^{p(l)} = \{e_1^{p(l)}, \dots, e_N^{p(l)}\}$ on the basis of the observations $\mathbf{es}^{X^c} = \{\mathbf{es}_1^{X^c}, \dots, \mathbf{es}_N^{X^c}\}$ from this parametric model.

We then choose the vector \mathbf{a} to minimize

$$\tilde{M}[\mathbf{a}; L, \mathbf{e}^{p^*}, \tilde{F}(\cdot | \mathbf{es}^{X^c})] = \frac{1}{L} \sum_{l=1}^L I[\mathbf{a}; \mathbf{es}^{X^c}, \mathbf{e}^{p(l)}, \mathbf{e}^{p^*}, \tilde{F}(\cdot | \mathbf{es}^{X^c})],$$

where

$$I[\mathbf{a}; \mathbf{es}^{X^c}, \mathbf{e}^{p(l)}, \mathbf{e}^{p^*}, \tilde{F}(\cdot | \mathbf{es}^{X^c})] = \frac{\Delta}{N} \sum_{j=1}^M \sum_{i=1}^N [\hat{F}(e_j^{p^*} | \mathbf{es}^{X^c} i, \mathbf{a}) - \tilde{F}(e_j^{p^*} | \mathbf{es}_i^{X^c})]^2$$

and \mathbf{e}^{p^*} is a vector of M evenly spaced values over the sample space of e_i^p , with $e_{j+1}^{p^*} - e_j^{p^*} = \Delta$. The term $\hat{F}(\cdot | \mathbf{es}_i^{X^c}, \mathbf{a})$ is our nonparametric estimate of the conditional distribution of e^p given \mathbf{es}^{X^c} , and $\tilde{F}(\cdot | \mathbf{es}_i^{X^c})$ is the parametric estimate of the condition distribution.

³¹ We also employed the cross-validation method set forth in Li and Racine (2008) to compute the bandwidth parameters. However, that method took an order of magnitude longer to run, because of the size of our data, without the resulting values of \mathbf{a} being very far from those estimated using the method of Hall, Wolff, and Yao (1999).

Given \mathbf{a}^{Opt} , the optimized value of \mathbf{a} , we can compute $\hat{F}(e^p | \mathbf{e}\mathbf{s}^{X^c}, \mathbf{a}^{\text{Opt}})$ for any values of e^p and $\mathbf{e}\mathbf{s}^{X^c}$. The process for computing the optimal value of \mathbf{a} described above can be automated, given a sample of $(e_i^p, \mathbf{X}_i^{c'})'$ for $i = 1, 2, \dots, N$, as each step of the process has a well-defined termination rule.

As more data become available, the process of updating our estimator of $F(y | \mathbf{X})$ involves first recomputing the residuals e_i^p and $\mathbf{e}_i^{X^c}$ and constructing $\mathbf{e}\mathbf{s}_i^{X^c}$ for the new sample. An updated value of \mathbf{a} would need to be computed. However, this would be simplified by the fact that the value from the previous sample can be used as a starting value for the procedure.

To compute the value of $\hat{F}(y | \mathbf{X})$ for a shipment with price p and characteristics \mathbf{X}_*^d and \mathbf{X}_*^c , we first use the value of \mathbf{X}_*^c and the estimated coefficients from the regressions of p on \mathbf{X}^d and elements of \mathbf{X}^c on \mathbf{X}^d for the sample of reasonable shipments to compute e_*^p , the difference between p_* and the predicted value given \mathbf{X}_*^d , and \mathbf{e}^{X^c} , the element-by-element difference between \mathbf{X}_*^c and the predicted value given \mathbf{X}_*^d . We then scale \mathbf{e}^{X^c} using standard deviations from the reasonable-price sample to compute $\mathbf{e}\mathbf{s}^{X^c} = \mathbf{S}^{-1} \mathbf{e}^{X^c}$. Given the value of $\hat{F}(e_*^p | \mathbf{e}\mathbf{s}^{X^c})$, this shipment price is classified as unreasonable for α , using the procedure described in Section 4.³²

References

- Akaike, H. 1973. Information Theory and an Extension of the Maximum Likelihood Principle. Pp. 267–81 in *Proceedings of the 2nd International Symposium on Information Theory*, edited by B. N. Petrov and F. Csáki. Budapest: Akademiai Kiado.
- Barnekov, C. C., and A. N. Kleit. 1990. The Efficiency Effects of Railroad Deregulation in the United States. *International Journal of Transport Economics* 17:21–36.
- Bashtannyk, David M., and Rob J. Hyndman. 2001. Bandwidth Selection for Kernel Conditional Density Estimation. *Computational Statistics and Data Analysis* 36:279–98.
- Berndt, Ernst R., Ann F. Friedlaender, Judy Shaw-Er Wang Chiang, and Christopher A. Velluro. 1993. Cost Effects of Mergers and Deregulation in the U.S. Rail Industry. *Journal of Productivity Analysis* 4:127–44.
- Bitzan, John D., and Theodore E. Keeler. 2003. Productivity Growth and Some of Its Determinants in the Deregulated U.S. Railroad Industry. *Southern Economic Journal* 70:232–53.
- Bitzan, John D., and Wesley W. Wilson. 2007. Industry Costs and Consolidation: Efficiency Gains and Mergers in the U.S. Railroad Industry. *Review of Industrial Organization* 30:81–105.
- Boyer, Kenneth D. 1987. The Cost of Price Regulation: Lessons from Railroad Deregulation. *RAND Journal of Economics* 18:408–16.
- Burton, Mark L. 1993. Railroad Deregulation, Carrier Behavior, and Shipper Response: A Disaggregated Analysis. *Journal of Regulatory Economics* 5:417–34.

³² We also experimented with the Li and Racine (2008) procedure for nonparametric conditional cumulative distribution function estimation with mixed categorical and continuous data and obtained similar results. This procedure is more computationally intensive and requires more user-selected parameters, which would make it more challenging to implement in a regulatory process than our simplified procedure.

- Burton, Mark, and Wesley W. Wilson. 2006. Network Pricing: Service Differentials, Scale Economies, and Vertical Exclusion in Railroad Markets. *Journal of Transport Economics and Policy* 40:255–77.
- Demsetz, Harold. 1968. Why Regulate Utilities? *Journal of Law and Economics* 11:55–65.
- . 1969. Information and Efficiency: Another Viewpoint. *Journal of Law and Economics* 12:1–22.
- Dennis, Scott M. 2001. Changes in Railroad Rates since the Staggers Act. *Transportation Research Part E: Logistics and Transportation Review* 37:55–69.
- Eaton, Jade Alice, and Judith A. Center. 1985. A Tale of Two Markets: The ICC's Use of Product and Geographic Competition in the Assessment of Rail Market Dominance. *Transportation Practitioners Journal* 53:16–35.
- Ellig, Jerry. 2002. Railroad Deregulation and Consumer Welfare. *Journal of Regulatory Economics* 21:143–67.
- Gallamore, Robert E., and John R. Meyer. 2014. *American Railroads: Decline and Renaissance in the Twentieth Century*. Cambridge, MA: Harvard University Press.
- Gaskins, Darius W., Jr. 2008. Regulation of Freight Railroads in the Modern Era: 1970–2010. *Review of Network Economics* 7:561–72.
- Government Accountability Office. 2006. *Freight Railroads: Industry Health Has Improved, but Concerns about Competition and Capacity Should Be Addressed*. Washington, DC: Government Accountability Office.
- Grimm, Curtis, and Clifford Winston. 2000. Competition in the Deregulated Railroad Industry: Sources, Effects, and Policy Issues. Pp. 41–71 in *Deregulation of Network Industries: What's Next*, edited by Sam Peltzman and Clifford Winston. Washington, DC: Brookings Institution Press.
- Grimm, Curtis M., Clifford Winston, and Carol A. Evans. 1992. Foreclosure of Railroad Markets: A Test of Chicago Leverage Theory. *Journal of Law and Economics* 35:295–310.
- Hall, Peter, Rodney C. L. Wolff, and Qiwei Yao. 1999. Methods for Estimating a Conditional Distribution Function. *Journal of the American Statistical Association* 94:154–63.
- Keeler, Theodore E. 1983. *Railroads, Freight, and Public Policy*. Washington, DC: Brookings Institution.
- Laffont, Jean-Jacques, and Jean Tirole. 1990a. The Regulation of Multiproduct Firms: Part I: Theory. *Journal of Public Economics* 43:1–36.
- . 1990b. The Regulation of Multiproduct Firms: Part II: Applications to Competitive Environments and Policy Analysis. *Journal of Public Economics* 43:37–66.
- Li, Qi, and Jeffrey Scott Racine. 2007. *Nonparametric Econometrics: Theory and Practice*. Princeton, NJ: Princeton University Press.
- . 2008. Nonparametric Estimation of Conditional CDF and Quantile Functions with Mixed Categorical and Continuous Data. *Journal of Business and Economic Statistics* 26:423–34.
- MacDonald, James M. 1987. Competition and Rail Rates for the Shipment of Corn, Soybeans, and Wheat. *RAND Journal of Economics* 18:151–63.
- . 1989. Railroad Deregulation, Innovation, and Competition: Effects of the Staggers Act on Grain Transportation. *Journal of Law and Economics* 32:63–95.
- MacDonald, James M., and Linda C. Cavalluzzo. 1996. Railroad Deregulation: Pricing Reforms, Shipper Responses, and the Effects on Labor. *Industrial and Labor Relations Review* 50:80–91.
- McFarland, Henry. 1989. The Effects of United States Railroad Deregulation on Shippers, Labor, and Capital. *Journal of Regulatory Economics* 1:259–70.

- Nadaraya, E. A. 1964. On Estimating Regression. *Theory of Probability and Its Applications* 9:141–42.
- Pagan, Adrian, and Aman Ullah. 1999. *Nonparametric Econometrics*. Cambridge: Cambridge University Press.
- Pittman, Russell. 2010. Against the Stand-Alone-Cost Test in U.S. Freight Rail Regulation. *Journal of Regulatory Economics* 38:313–26.
- Railroad-Shipper Transportation Advisory Council. 2011. Position Paper on the Uniform Rail Costing System. November 22. Railroad-Shipper Transportation Advisory Council, Washington, DC. <https://www.stb.gov/wp-content/uploads/files/docs/railroadShipperTransportation/RSTAC%20URCS%20White%20Paper%20on%20URCS%20November%2022.pdf>.
- Rate Reform Task Force. 2019. *Report to the Surface Transportation Board*. April 25. Washington, DC: Rate Reform Task Force. <https://prod.stb.gov/wp-content/uploads/Rate-Reform-Task-Force-Report-April-2019.pdf>.
- Schmalensee, Richard L., and Wesley W. Wilson. 2016. Modernizing U.S. Freight Rail Regulation. *Review of Industrial Organization* 49:133–59.
- Schmidt, Stephen. 2001. Market Structure and Market Outcomes in Deregulated Rail Freight Markets. *International Journal of Industrial Organization* 19:99–131.
- Stone, Richard D. 1991. *The Interstate Commerce Commission and the Railroad Industry: A History of Regulatory Policy*. New York: Praeger.
- TRB-NAS (Transportation Research Board and National Academies of Sciences, Engineering, and Medicine). 2015. *Modernizing Freight Rail Regulation*. Special Report No. 318. Washington, DC: National Academies Press.
- Velluro, Christopher A., Ernst R. Berndt, Ann F. Friedlaender, Judy Shaw-Er Wang Chiang, and Mark H. Showalter. 1992. Deregulation, Mergers, and Cost Savings in Class 1 U.S. Railroads, 1974–1986. *Journal of Economics and Management Strategy* 1:339–69.
- Watson, Geoffrey S. 1964. Smooth Regression Analysis. *Sankhyā: The Indian Journal of Statistics, Series A* 26:359–72.
- Wilson, Wesley W. 1994. Market-Specific Effects of Rail Deregulation. *Journal of Industrial Economics* 42:1–22.
- . 1996. Legislated Market Dominance in Railroad Markets. *Research in Transportation Economics* 4:49–67.
- . 1997. Cost Savings and Productivity in the Railroad Industry. *Journal of Regulatory Economics* 11:21–40.
- Wilson, Wesley W., and Frank A. Wolak. 2016. Freight Rail Costing and Regulation: The Uniform Rail Costing System. *Review of Industrial Organization* 49:229–61.
- Winston, Clifford, Thomas M. Corsi, Curtis M. Grimm, and Carol A. Evans. 1990. *The Economic Effects of Surface Freight Deregulation*. Washington, DC: Brookings Institution Press.